Physica D 51 (1991) 177–188
North-Holland

Homoclinic chaos for ray optics in a fiber

Darryl D. Holm and Gregor Kovačič

Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, MS B284, Los Alamos, NM 87545, USA

In the geometrical optics approximation, stable and unstable manifolds of periodic orbits, invariant tori, and hyperbolic invariant manifolds are shown to exist and produce trapping of bundles of light rays near the axis of a translation-invariant, axisymmetric optical fiber whose squared refractive index is a parabolic function of squared radius. Periodic symmetry-breaking perturbations in the refractive index are shown to destroy this ray trapping and to produce homoclinic tangles, through which nearby trapped rays may escape, and untrapped rays may become at least temporarily trapped. Melnikov's technique is used to prove that the perturbations cause the stable and unstable manifolds of the unperturbed periodic orbits, invariant tori, and hyperbolic invariant manifolds to develop transverse intersections and therefore, to form homoclinic tangles. These tangles imply either homoclinic chaos by the Smale horseshoe mechanism and the Poincaré–Birkhoff–Smale theorem, or Arnol'd diffusion. In both situations, lobe dynamics will dominate phase space transport, seen here as a flux of (initially) untrapped rays passing through the trapping region.

1. Introduction

In geometrical optics, the ray path is determined by Fermat's principle of least optical length,

$$\delta \int n \, \mathrm{d}s = 0,$$

where $n(x, y, z)$ is the index of refraction at the spatial point (x, y, z) and $\mathrm{d}s$ is the element of arc length along the ray path through that point. Choosing coordinates so that the z-axis coincides with the optical axis of the fiber, gives

$$\begin{aligned} \mathrm{d}s &= \left[(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2\right]^{1/2} \\ &= \left[1 + \dot{x}^2 + \dot{y}^2\right]^{1/2} \mathrm{d}z, \end{aligned}$$

with $\dot{x} = \mathrm{d}x/\mathrm{d}z$ and $\dot{y} = \mathrm{d}y/\mathrm{d}z$. Thus, Fermat's principle can be written in Lagrangian form, with z playing the role of time,

$$\delta \int L \, \mathrm{d}z = 0,$$

and where the optical Lagrangian is given by

$$L = n(x, y, z)\left[1 + \dot{x}^2 + \dot{y}^2\right]^{1/2} = \frac{n}{\gamma}.$$

Here, $\gamma = \mathrm{d}z/\mathrm{d}s$ and variations are taken in x and y at constant z. The Euler–Lagrange equations determine the ray path, $\boldsymbol{q}(z)$, in two-component vector form according to

$$\frac{\mathrm{d}}{\mathrm{d}z}\frac{\partial L}{\partial \dot{\boldsymbol{q}}} = \frac{\partial L}{\partial \boldsymbol{q}}, \quad \boldsymbol{q} = (x, y),$$

with

$$L = n(\boldsymbol{q}, z)\left[1 + |\dot{\boldsymbol{q}}|^2\right]^{1/2} = \frac{n}{\gamma}.$$

Specifically, the vector Euler–Lagrange equation of the light rays is

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}s}\right) = \gamma\frac{\mathrm{d}}{\mathrm{d}z}\left(n\gamma\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}z}\right) = \frac{\partial n}{\partial \boldsymbol{q}}.$$

The momentum $\boldsymbol{p}$ canonically conjugate to the ray path position $\boldsymbol{q}$ in an "image plane", or

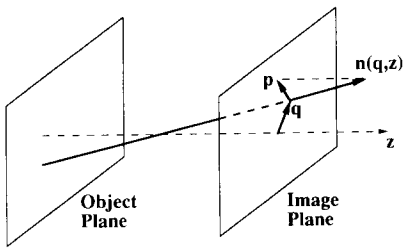

Fig. 1. The momentum $\boldsymbol{p}$ canonically conjugate to the coordinate $\boldsymbol{q}$ on the image plane at z is the projection onto the plane of the ray vector $\boldsymbol{n}(\boldsymbol{q}, z)$ through the point $\boldsymbol{q}$. That is, $|\boldsymbol{p}| = n(\boldsymbol{q}, z)\cos\theta$.

"image screen", at a fixed value of z is given by

$$\boldsymbol{p} = \frac{\partial L}{\partial \dot{\boldsymbol{q}}} = n\gamma\dot{\boldsymbol{q}},$$

and satisfies

$$|\boldsymbol{p}|^2 = n^2(1-\gamma^2).$$

Defining $\sin\theta = \mathrm{d}z/\mathrm{d}s = \gamma$ leads to

$$|\boldsymbol{p}| = n\cos\theta,$$

and gives the geometrical picture of the ray path shown in fig. 1. Along the optical axis (the z-axis) each image plane normal to the axis is pierced at a point $\boldsymbol{q} = (x, y)$ by a vector of magnitude $n(\boldsymbol{q}, z)$ tangent to the ray path and that makes an angle θ to the plane. The projection of this vector onto the image plane is the canonical momentum $\boldsymbol{p}$. This picture of the ray paths captures all but the rays of grazing incidence to the image planes. Such grazing rays are ignored in what follows.

Passing now via the usual Legendre transformation from the Lagrangian to the Hamiltonian description gives

$$H = \boldsymbol{p}\cdot\dot{\boldsymbol{q}} - L = n\gamma|\dot{\boldsymbol{q}}|^2 - \frac{n}{\gamma}$$
$$= -n\gamma = -\sqrt{n^2(\boldsymbol{q}, z) - |\boldsymbol{p}|^2}\,.$$

Thus, in the geometrical picture, the component of the ray-path tangent vector along the optical axis is (minus) the Hamiltonian, i.e. $n\sin\theta = -H$, also shown in fig. 1. The phase space description of the ray path now follows from Hamilton's equations,

$$\dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}} = -\frac{1}{H}\boldsymbol{p}, \quad \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}} = -\frac{1}{2H}\frac{\partial n^2}{\partial \boldsymbol{q}}.$$

Remark. If $n = n(\boldsymbol{q})$, so that the medium is translation invariant along the optical axis, z, then $H = -n\sin\theta$ is conserved. (Conservation of H at an interface is Snell's law.) For translation-invariant media, the vector ray-path equation simplifies to

$$\ddot{\boldsymbol{q}} = \frac{1}{2H^2}\frac{\partial n^2}{\partial \boldsymbol{q}}.$$

Thus, in this case geometrical ray tracing reduces to "Newtonian dynamics" in z, with potential $-n^2(\boldsymbol{q})$ and with "time" rescaled along each path by the value of $\sqrt{2}\,H$ determined from the initial conditions for each ray.

2. Axisymmetric, translation-invariant media

Axisymmetric, translation-invariant media, in which the index of refraction is a function of the radius alone, are of considerable theoretical interest. Axisymmetry implies an additional constant of motion and, hence, reduction of the Hamiltonian system for the light rays to phase plane analysis. The resulting ray paths describe, in a certain sense, perfect optical instruments [6]. For such media, the index of refraction satisfies

$$n(\boldsymbol{q}, z) = n(r), \quad r = |\boldsymbol{q}|.$$

Passing to polar coordinates (r, φ) with $\boldsymbol{q} = (x, y) = r(\cos\varphi, \sin\varphi)$ leads in the usual way to

$$|\boldsymbol{p}|^2 = p_r^2 + \frac{p_\varphi^2}{r^2}.$$

Consequently, the optical Hamiltonian,

$$H = -\sqrt{n^2(r) - p_r^2 - \frac{p_\varphi^2}{r^2}},$$

is independent of the azimuthal angle φ; so the canonically conjugate "angular momentum" p_φ is conserved.

Using the relation $\boldsymbol{p}\cdot\boldsymbol{q} = rp_r$ leads to an interpretation of p_φ in terms of the image-screen phase space variables $\boldsymbol{p}$ and $\boldsymbol{q}$. Namely,

$$|\boldsymbol{p}\times\boldsymbol{q}|^2 = |\boldsymbol{p}|^2|\boldsymbol{q}|^2 - (\boldsymbol{p}\cdot\boldsymbol{q})^2 = p_\varphi^2.$$

The conserved quantity $p_\varphi = \boldsymbol{p}\times\boldsymbol{q} = yp_x - xp_y$ is called the *skewness function*, or the Petzval invariant for axisymmetric media [14]. Vanishing of p_φ occurs for *meridional* rays, for which $\boldsymbol{p}$ and $\boldsymbol{q}$ are collinear in the image plane. On the other hand, p_φ takes its maximum value for *sagittal* rays, for which $\boldsymbol{p}\cdot\boldsymbol{q} = 0$, so that $\boldsymbol{p}$ and $\boldsymbol{q}$ are orthogonal in the image plane. (See ref. [2], pp. 151 and 190, for further discussion of the properties of meridional, sagittal, and other special rays.)

Hamilton's equations for axisymmetric, translation-invariant media are expressible as

$$\dot r = \frac{\partial H}{\partial p_r} = -\frac{1}{H}p_r,$$

$$\dot p_r = -\frac{\partial H}{\partial r} = -\frac{1}{2H}\frac{\mathrm{d}}{\mathrm{d}r}\left(n^2(r) - \frac{p_\varphi^2}{r^2}\right),$$

$$\dot\varphi = \frac{\partial H}{\partial p_\varphi} = -\frac{p_\varphi}{Hr^2},$$

$$\dot p_\varphi = -\frac{\partial H}{\partial \varphi} = 0.$$

We shall solve these equations for the case of an optical fiber with radially graded index of refraction. We choose a radial profile of the fiber's refractive index in the following form:

$$n^2(r) = \lambda^2 + (\mu - \nu r^2)^2,$$

where λ^2, μ and ν are positive constants. Accordingly, the fiber's refractive index is a parabola in the squared radius with its minimum at $r^2 = \mu/\nu$. In this case, the level surfaces of the Hamiltonian form a family of curves in the (r, p_r) phase plane given by

$$p_r^2 = \lambda^2 + (\mu - \nu r^2)^2 - \frac{p_\varphi^2}{r^2} - H^2.$$

These curves allow positive H^2 provided

$$\lambda^2 + (\mu - \nu r^2)^2 - \frac{p_\varphi^2}{r^2} - p_r^2 \geq 0.$$

This is the physical domain of r and p_r. This domain in the (r, p_r) phase plane is shown as the unshaded regions of fig. 2, for $p_\varphi = 0$, and of fig. 3, for $p_\varphi \neq 0$. Fig. 2 shows the pair of heteroclinic orbits appearing in the (r, p_r) phase plane for $p_\varphi = 0$, while fig. 3 shows the homoclinic orbit appearing for $p_\varphi \neq 0$. In both cases, the stable and unstable manifolds of the hyperbolic equilibrium points have branches going off to infinity. A

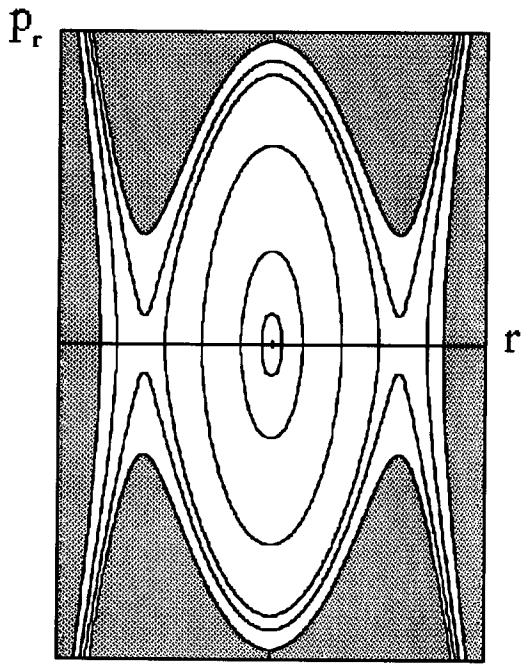

Fig. 2. The unshaded region is the physical domain of (r, p_r) phase space for $p_\varphi = 0$.

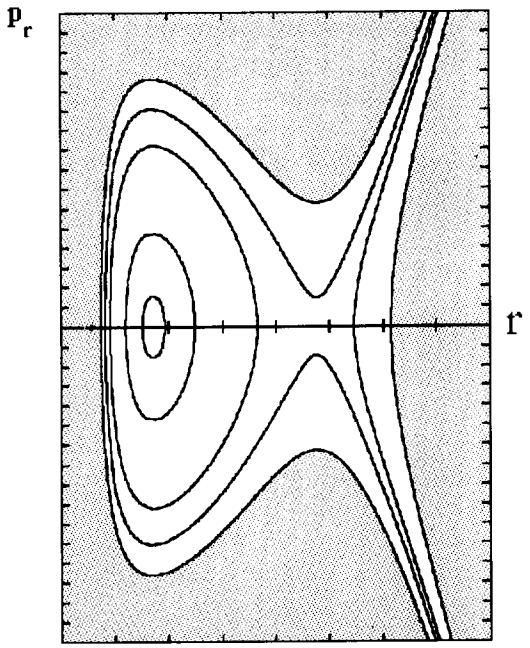

Fig. 3. The unshaded region is the physical domain of (r, p_r) phase space for $p_\varphi \neq 0$.

level surface of the Hamiltonian H in (r, p_r, p_φ) is shown in fig. 4. The intersections of the level surfaces of H with a (nonzero) level surface of p_φ produce the phase space portrait shown in fig. 3.

Hamilton's equations for (r, p_r) which result for this choice of radially graded refractive index are equivalent to the dynamics of an ideal (inverted) Duffing oscillator in a rotating frame (up to the rescaling of time by the energy on each orbit). Namely,

$$\dot{r} = -\frac{1}{H} p_r,$$

$$\dot{p}_r = -\frac{1}{H}\left(-2\nu r(\mu - \nu r^2) + \frac{p_\varphi^2}{r^3}\right).$$

The equilibria of these equations occur at the points where the right-hand sides vanish, along the r-axis (i.e. $p_r = 0$) at the values $r = \bar{r}(p_\varphi; \mu, \lambda)$ satisfying the following cubic relation in r^2:

$$-2\nu r^2(\mu - \nu r^2) + \frac{p_\varphi^2}{r^2} = 0.$$

The level surfaces of H^2 satisfy this equilibrium condition when

$$\lambda^2 + (\mu - \nu r^2)^2 - H^2 = 2\nu r^2(\mu - \nu r^2).$$

Thus, the equilibrium points occur in pairs in r^2 at the roots of this quadratic, for $\lambda^2 < \frac{1}{3}(\mu^2 + H^2)$,

$$r^2 = \frac{1}{3\nu}\left(2\mu \pm \sqrt{\mu^2 - 3\lambda^2 + H^2}\right).$$

At the value $\lambda^2 = \frac{1}{3}(\mu^2 + H^2)$ the homoclinic loop closes to a cusp. For $\lambda^2 > \frac{1}{3}(\mu^2 + H^2)$, no equilibrium occurs in the physical domain. In what follows, we shall assume that $\lambda^2 < \frac{1}{3}(\mu^2 + H^2)$.

An equilibrium point is stable, provided at that point

$$\frac{d^2}{dr^2}\left(n^2 - \frac{p_\varphi^2}{r^2}\right) < 0.$$

In each pair of roots in r^2 of the equilibrium condition, the larger one is unstable, while the smaller one is stable. The unstable root is connected to itself by a homoclinic orbit encircling the stable root. (The two other orbits connected to the unstable root are branches connected to plus and minus infinity.) Initial values lying within the homoclinic loop are trapped forever within it on periodic orbits, while those lying initially outside the loop escape to infinity.

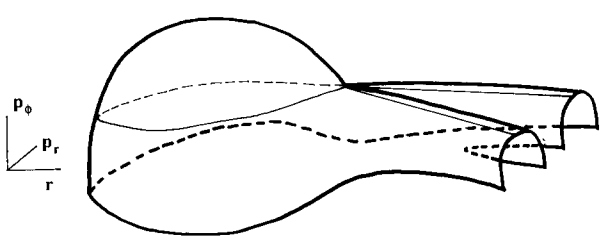

Fig. 4. Level surface of $H_0(r, p_r, p_\varphi)$ away from $p_\varphi = 0$.

The homoclinic orbits (for $p_\varphi \neq 0$) and heteroclinic orbits (for $p_\varphi = 0$) are obtained by solving

$$\dot{r}^2 = (\mu - \nu r^2)^2 - \frac{p_\varphi^2}{r^2} - (\bar{H}^2 - \lambda^2),$$

where $\bar{H}$ is the value of the Hamiltonian at the unstable equilibrium point $\bar{r}$.

In the (degenerate) case, when $p_\varphi = 0$ (i.e. for meridional rays) there is an unstable equilibrium at $r^2 = \mu/\nu$, and a stable one at $r^2 = 0$. In this case, the two unstable equilibria are connected to each other by two *heteroclinic* orbits. The value of the squared Hamiltonian $\bar{H}^2$ for this orbit is λ^2. Hence, the heteroclinic orbits given by solving the previous equation in this case are

$$r^{(0)}(z) = \sqrt{\mu/\nu}\,\tanh\left(\pm z\sqrt{\mu\nu}\right),$$
$$p_r^{(0)}(z) = \pm\mu\,\mathrm{sech}^2\left(\pm z\sqrt{\mu\nu}\right),$$

where $\pm$ distinguishes the two branches connecting the equilibrium points.

The explicit expressions for the homoclinic and heteroclinic orbits are not actually needed in what follows. Only the reflection symmetry of these orbits under $z \to -z$ is needed, in order to determine the effects of perturbations of the refractive index on the stable and unstable manifolds of the hyperbolic equilibrium points. In particular, when $p_\varphi = 0, r^{(0)}(z)$ is odd, while $p_r^{(0)}(z)$ is even, as seen from the above equations and fig. 2. When $p_\varphi \neq 0, r^{(0)}(z)$ is even, while $p_r^{(0)}(z)$ is odd, as seen from fig. 3. From the expression $\dot{\varphi} = -p_\varphi/Hr^2$ we also see that $\dot{\varphi}$ is even for $p_\varphi \neq 0$, so the solution for φ may be written as $\varphi = \varphi^{(0)}(z) + \varphi_0$, where $\varphi^{(0)}(z)$ is odd, and φ_0 is an arbitrary constant. (The function $\varphi^{(0)}(z)$ is constant for $p_\varphi = 0$.)

3. Effects of perturbations of the refractive index

We consider the following axial and azimuthal perturbations of the refractive index:

$$n^2(r) = n_0^2(r) + \varepsilon n_1^2(r, z, \varphi).$$

Expanding the Hamiltonian

$$H = -\sqrt{n_0^2(r) - |\boldsymbol{p}|^2 + \varepsilon n_1^2(r, z, \varphi)}$$

in powers of ε, gives $H_\varepsilon = H_0 + \varepsilon H_1$ at linear order, where

$$H_0 = \sqrt{n_0^2(r) - |\boldsymbol{p}|^2}, \quad H_1 = \frac{1}{2H_0} n_1^2(r, z, \varphi).$$

This expansion is valid provided $H_0 \gg \varepsilon$; that is, away from grazing incidence to the image screen.

To linear order in ε, the refractive index perturbations appear in Hamilton's equations as

$$\dot{r} = \frac{\partial H_\varepsilon}{\partial p_r} = -\frac{1}{H_0} p_r + \varepsilon \frac{n_1^2}{2H_0^3} p_r,$$

$$\dot{p}_r = -\frac{\partial H_\varepsilon}{\partial r}$$
$$= -\frac{1}{2H_0}\frac{\mathrm{d}}{\mathrm{d}r}\left(n_0^2(r) - \frac{p_\varphi^2}{r^2}\right)$$
$$+ \varepsilon\left[\frac{n_1^2}{4H_0^3}\frac{\mathrm{d}}{\mathrm{d}r}\left(n_0^2(r) - \frac{p_\varphi^2}{r^2}\right) - \frac{1}{2H_0}\frac{\partial n_1^2}{\partial r}\right],$$

$$\dot{\varphi} = \frac{\partial H_\varepsilon}{\partial p_\varphi} = -\frac{p_\varphi}{H_0 r^2} + \varepsilon \frac{n_1^2}{2H_0^3}\frac{p_\varphi}{r^2},$$

$$\dot{p}_\varphi = -\frac{\partial H_\varepsilon}{\partial \varphi} = -\frac{\varepsilon}{2H_0}\frac{\partial n_1^2}{\partial \varphi}.$$

Explicitly, we consider axially and azimuthally *periodic* perturbations of the refractive index in the form

$$n_1^2 = \alpha r \sin(\omega z) + \beta \sin(m\varphi),$$

where α, β and ω are real constants, and m is an integer. There are three cases: (I) purely axial perturbations, $\alpha \neq 0$, $\beta = 0$; (II) purely azimuthal ones, $\alpha = 0$, $\beta \neq 0$; and (III) combined perturbations, $\alpha \neq 0$, $\beta \neq 0$.

3.1. Case (I). Purely axial perturbations, $n_1^2 = \alpha r \sin(\omega z)$

In this case, the four-dimensional phase space $(r, p_r, \varphi, p_\varphi)$ is foliated by preserved level surfaces of p_φ. Moreover, the φ-dynamics still separates out as a quadrature. Therefore, the perturbed dynamics reduces to motion in r, p_r and z, along level surfaces of p_φ.

The periodic perturbation induces a Poincaré map of the (r, p_r) plane in the form

$$(r, p_r)(z) \to (r, p_r)(z + 2\pi/\omega).$$

Suppose the perturbation causes transverse intersections of stable and unstable manifolds of the hyperbolic equilibria under this Poincaré map. Then, according to the Poincaré–Birkhoff–Smale theorem, these intersections imply homoclinic chaos via the construction of a horseshoe map from an iterate of the Poincaré map. That is, under some number of iterates of the Poincaré map, a nearly rectangular region of phase points, bounded by segments of the stable and unstable manifolds of the hyperbolic equilibrium and lying initially near the hyperbolic equilibrium, will become folded, stretched, contracted, and eventually mapped back over itself in the shape of a horseshoe. This horseshoe map is the underlying mechanism for chaos. Forward and backward iteration of the original rectangular region under the horseshoe map creates an invariant Cantor set structure within the area circumscribed by the original rectangular region. This invariant Cantor set can be shown to contain countably many unstable periodic motions, and uncountably many unstable nonperiodic motions. (See refs. [3, 11, 12] for discussions of the methods of proof of these statements and further descriptions of this type of homoclinic chaos.)

In order to prove the existence of transverse intersections of stable and unstable manifolds of hyperbolic equilibria under this Poincaré map, we use Melnikov's theorem [7]. This theorem states that simple zeros in the signed distance between the stable and unstable manifolds of a homoclinic point at *linear* order in perturbation theory are sufficient to imply transverse intersections of these manifolds under the full nonlinear dynamics. In the present case, these intersections take place in the Poincaré map of the (r, p_r) plane; so this distance takes the well-known scalar form defined by [3, 12]

$$M(z_0) = \int_{-\infty}^{\infty} \{H_0, H_1\}\big(r^{(0)}(z), p_r^{(0)}(z), z + z_0\big)\,dz,$$

where $\{H_0, H_1\}$ is the canonical Poisson bracket between H_0 and H_1, and the integral is taken along an unperturbed homoclinic orbit $(r^{(0)}(z), p_r^{(0)}(z))$, with phase delay z_0 in the *explicit* time dependence. The Hamiltonian H_0 is constant along the unperturbed homoclinic orbit. Recall that for $p_\varphi \neq 0$, $r^{(0)}(z)$ is an even function and $p_r^{(0)}(z)$ is odd. (When $p_\varphi = 0$, the function $r^{(0)}(z)$ is odd and $p_r^{(0)}(z)$ is even.) For axial perturbations we have

$$\{H_0, H_1\}\big(r^{(0)}(z), p_r^{(0)}(z), z + z_0\big) = -\frac{\alpha}{2H_0^2} p_r^{(0)}(z) \sin\big[\omega(z + z_0)\big].$$

Hence, expanding $\sin \omega(z + z_0)$ and integrating gives (for $p_\varphi \neq 0$)

$$M(z_0) = -\frac{\alpha A}{2H_0^2} \cos(\omega z_0),$$

with

$$A = \int_{-\infty}^{\infty} p_r^{(0)}(z) \sin(\omega z)\,dz.$$

Since the integrand is even, $A \neq 0$ and the Melnikov function $M(z_0)$ has simple zeros at half-odd-integer multiples of $\omega/2\pi$ for $p_\varphi \neq 0$. Consequently, the perturbation causes transverse intersections in the Poincaré map of stable and unstable manifolds of the hyperbolic equilibrium. Hence, horseshoe chaos appears in the Poincaré map under this perturbation. (When $p_\varphi = 0$, similar expressions for M and A occur, but with sin

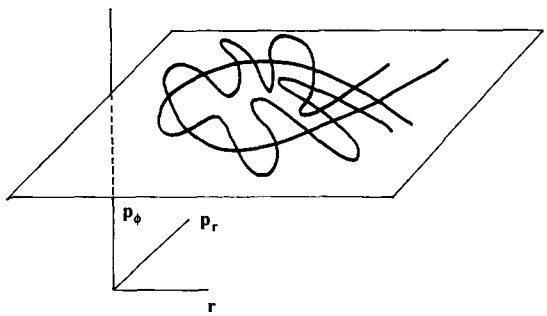

Fig. 5. Poincaré map showing the homoclinic tangle on a level surface of p_φ created by a purely axial, periodic perturbation.

and cos exchanged; so the zeros of $M(z_0)$ occur at *integer* multiples of $\omega/2\pi$ in this case.) See fig. 5 for a sketch of the homoclinic tangle produced in the Poincaré map by these purely axial, periodic perturbations.

The results of Rom-Kedar et al. [9, 10] show that phase space transport takes place within the homoclinic tangle on each level surface of the skewness p_φ as a result of the purely axial perturbations. This phase space transport is chaotic (in the sense of extreme sensitivity to initial conditions) and may cause trapped rays starting near the unperturbed homoclinic orbit to become untrapped, and vice versa. The stable and unstable manifolds of the unperturbed hyperbolic equilibrium on each level surface of p_φ both form a homoclinic orbit and have branches going off to infinity. For this topology, the results of Rom-Kedar et al. [9, 10] show that no initially untrapped rays may become permanently trapped under the perturbation. Hence, the axial perturbation induces a flux of initially untrapped rays through the trapping region on each level surface of the skewness p_φ. This flux may be computed numerically using the methods of Rom-Kedar et al.

3.2. *Case (II). Purely azimuthal perturbations, $n_1^2 = \beta \sin(m\varphi)$*

In this case, the unperturbed dynamics takes place in the four-dimensional phase space $(r, p_r, \varphi, p_\varphi)$ on three-dimensional level surfaces of H_0. (This case fits into the "System III" framework of Wiggins [11].) The unstable equilibrium in the (r, p_r) phase plane is a periodic orbit S_0 in the full phase space. Each periodic orbit S_0 lies on the intersection of a level surface of H_0 with the two-dimensional, normally hyperbolic manifold $\mathcal{M}_0$ coordinatized by (φ, p_φ). (A level surface of H_0 in (r, p_r, p_φ) coordinates in the region away from $p_\varphi = 0$ is shown in fig. 4. Intersections of such level surfaces of H_0 with the level surfaces of p_φ produce the phase space portrait shown in fig. 3.). By persistency of normally hyperbolic manifolds under perturbations, there exists a two-dimensional perturbed normally hyperbolic manifold $\mathcal{M}_\varepsilon$ close to the original $\mathcal{M}_0$ which is intersected transversely along periodic orbits S_ε by level surfaces of $H_\varepsilon = H_0 + \varepsilon H_1$. (See also ref. [4].) The stable and unstable manifolds of these orbits are each two-dimensional, and lie in the three-dimensional level surface of H_ε. Thus, to apply the Melnikov technique, we must seek conditions for these two-dimensional manifolds to intersect transversely in the three-dimensional level surface H_ε = constant. (The perturbed system is autonomous, but the perturbed φ-dynamics no longer separates out as a quadrature.)

The manifold $W(S_0)$ of orbits that are homoclinic to an orbit S_0 (periodic in $\varphi \in [0, 2\pi)$) at $r = \bar{r}(p_\varphi; \mu, \lambda, \nu)$, $p_r = 0$, and $p_\varphi = \text{const}$, is given by

$$H(r, p_r, p_\varphi) - \bar{H}(\bar{r}, 0, p_\varphi) = 0, \quad p_\varphi = \text{const}.$$

The distance between stable and unstable manifolds of the perturbed periodic orbits S_ε on $\mathcal{M}_\varepsilon$ is measured along the unperturbed normals to $W(S_0)$. Namely,

$$\boldsymbol{n}_1 = \nabla(H - \bar{H}), \quad \boldsymbol{n}_2 = \nabla p_\varphi,$$

where the gradient in the four-dimensional phase

space is given by

$$\nabla = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial p_r}, \frac{\partial}{\partial \varphi}, \frac{\partial}{\partial p_\varphi} \right).$$

In the next subsection, where axial and azimuthal perturbations are combined, we will discuss in more detail the actual mechanics of choosing points on the perturbed stable and unstable manifolds of the perturbed manifold $\mathcal{M}_\varepsilon$ and measuring the distance between them using the Melnikov method. For now, we only note that intersections of the stable and unstable manifolds of $\mathcal{M}_\varepsilon$ occur when the components of distance measured along both normal vectors vanish. In the present case, the relation $H_\varepsilon = \text{const}$ implies a functional dependence between these two components of distance, so both components vanish if either of them does. Thus, for convenience, we may choose to compute the distance between stable and unstable manifolds along just one component. The component we choose will be along the normal vector $\boldsymbol{n}_1$, since the integrals obtained for this choice are exponentially convergent. To first order in ε, this distance is proportional to the following integral along the unperturbed homoclinic orbit:

$$M_1 = \int_{-\infty}^{\infty} \boldsymbol{n}_1 \cdot \boldsymbol{g} \, \mathrm{d}z,$$

where $\boldsymbol{g}$ is the order ε part of the vector field. In this case,

$$\boldsymbol{n}_1 \cdot \boldsymbol{g} = \frac{p_\varphi}{2H_0^2} \left(\frac{1}{r^2} - \frac{1}{\bar{r}^2} \right) \frac{\partial n_1^2}{\partial \varphi}.$$

Now, $\partial n_1^2/\partial\varphi = m \cos\{m[\varphi^{(0)}(z) + \varphi_0]\}$, while the quantity $p_\varphi/2H_0^2$ is constant and the unperturbed radius $r^{(0)}(z)$ is an even function along the unperturbed homoclinic orbit for $p_\varphi \neq 0$. (Recall fig. 3.) Consequently,

$$M_1 = B \frac{m p_\varphi}{2H_0^2} \cos(m\varphi_0),$$

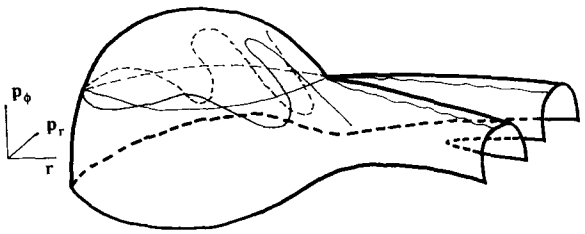

Fig. 6. Poincaré map showing the homoclinic tangle on a level surface of H_ε created by a purely azimuthal perturbation.

where

$$B = \int_{-\infty}^{\infty} \left(\frac{1}{r^2} - \frac{1}{\bar{r}^2} \right) \cos\left[m\varphi^{(0)}(z) \right] \mathrm{d}z.$$

Being the integral of an even function, $B \neq 0$, and the Melnikov function $M_1(\varphi_0)$ has simple zeros at $\varphi_0 = \pm\pi/2m$ for $p_\varphi \neq 0$, thereby implying transverse intersections of stable and unstable manifolds of the hyperbolic periodic orbits in $\mathcal{M}_\varepsilon$ under this perturbation.

A Poincaré section of the perturbed dynamics may be constructed by fixing a value $\varphi = \varphi_P$ and plotting recurrences of phase points in this three-dimensional surface. This three-dimensional Poincaré section has a two-dimensional intersection with the level surface of H_ε containing the perturbed periodic orbit. On this two-dimensional intersection, the periodic orbit appears as a point, and its stable and unstable manifolds appear as curves, which intersect transversely in a tangle. The homoclinic tangle produced by these intersections in the Poincaré section on a level surface of H_ε is sketched in fig. 6. The Poincaré–Birkhoff–Smale homoclinic chaos theorem now implies the existence of a horseshoe map for a sufficiently large iterate of the Poincaré map. (See ref. [11], discussion of "System III".)

Again, the presence of the homoclinic tangle implies chaotic phase space transport, which now takes place on each energy surface H_ε in the region of the tangle under purely azimuthal perturbations. This phase space transport may cause

trapped rays starting near the unperturbed homoclinic orbit to become untrapped, and vice versa. The stable and unstable manifolds of the unperturbed hyperbolic equilibrium on each level surface of energy H_0 form a homoclinic orbit, and have branches going off to infinity. For this topology, the results of Rom-Kedar et al. [9, 10] show that no initially untrapped rays may become permanently trapped under the perturbation. Hence, the azimuthal perturbation will induce a flux of initially untrapped rays through the trapping region on each level surface of the energy H_ε. Again, this flux could be computed numerically using the methods of Rom-Kedar et al.

3.3. Case (III). Combined perturbations, $n_1^2 = \alpha r \sin(\omega z) + \beta \sin(m\varphi)$

In this case of combined perturbations, the perturbed dynamics takes place in the five-dimensional space $(r, p_r, \varphi, p_\varphi, z)$. The unperturbed fixed points in the (r, p_r) plane now comprise a normally hyperbolic manifold $\mathcal{M}_0$, foliated by two-tori T_0 at $p_\varphi = \text{const}$, and parametrized by (φ, z). The manifold $\mathcal{M}_0$ has four-dimensional stable and unstable manifolds, $W_0^s(\mathcal{M}_0)$ and $W_0^u(\mathcal{M}_0)$, respectively, whose two branches coincide to form the unperturbed homoclinic manifold $W(\mathcal{M}_0)$. By persistency of normally hyperbolic manifolds under perturbations, $\mathcal{M}_0$ will deform to another normally hyperbolic manifold, $\mathcal{M}_\varepsilon$, which will also have four-dimensional stable and unstable manifolds, $W_\varepsilon^s(\mathcal{M}_\varepsilon)$ and $W_\varepsilon^u(\mathcal{M}_\varepsilon)$, respectively. Now the KAM theorem shows that most of the tori T_0 on $\mathcal{M}_0$ will survive as tori T_ε on $\mathcal{M}_\varepsilon$. (See, e.g. refs. [5, 11].) Moreover, their stable and unstable manifolds $W_\varepsilon^s(T_\varepsilon)$ and $W_\varepsilon^u(T_\varepsilon)$ will be three-dimensional ($T^2 \times \mathbb{R}^1$) and will be close to the unperturbed three-dimensional stable and unstable manifolds of the tori T_0 on $\mathcal{M}_0$. (Two branches of these unperturbed manifolds $W_0^s(T_0)$ and $W_0^u(T_0)$ coincide, to form the unperturbed homoclinic manifold $W(T_0)$.)

We are now in a position to discuss conditions for the occurrence of intersections of the perturbed stable and unstable manifolds of $\mathcal{M}_\varepsilon$. We shall use the unperturbed manifold $W(\mathcal{M}_0)$ to provide coordinates for the perturbed manifolds $W_\varepsilon^s(\mathcal{M}_\varepsilon)$ and $W_\varepsilon^u(\mathcal{M}_\varepsilon)$. The manifold $W(\mathcal{M}_0)$ is given by

$$H(r, p_r, p_\varphi) - \bar{H}(\bar{r}, 0, p_\varphi) = 0.$$

Consider the normal vector $\boldsymbol{n}_1$ at a point a on $W(\mathcal{M}_0)$. This normal vector cuts $W(\mathcal{M}_0)$ in precisely one point, namely a. Because transverse intersections persist under perturbations, the normal vector $\boldsymbol{n}_1$ will also cut both $W_\varepsilon^s(\mathcal{M}_\varepsilon)$ and $W_\varepsilon^u(\mathcal{M}_\varepsilon)$ in discrete points. One chooses to consider the *nearest* points (in terms of "time of flight", see refs. [11, 12]) to $\mathcal{M}_\varepsilon$ on each of $W_\varepsilon^s(\mathcal{M}_\varepsilon)$ and $W_\varepsilon^u(\mathcal{M}_\varepsilon)$, which are pierced by the normal vector $\boldsymbol{n}_1$. We denote these points by a_ε^s and a_ε^u, respectively, and measure the (signed) distance between them along $\boldsymbol{n}_1$ by the difference $a_\varepsilon^s - a_\varepsilon^u$. To first order in ε the distance along $\boldsymbol{n}_1$ is proportional to

$$\begin{aligned} M_1 &= \int_{-\infty}^{\infty} \boldsymbol{n}_1 \cdot \boldsymbol{g}\big(r^{(0)}(z), p_r^{(0)}(z), \varphi^{(0)}(z) + \varphi_0, \\ &\qquad z + z_0\big)\, dz \\ &= \frac{\alpha}{2H_0^2} A \cos(\omega z_0) + \frac{m\beta p_\varphi}{2H_0^2} B \cos(m\varphi_0), \end{aligned}$$

where the coefficients A and B are given as in cases (I) and (II), respectively. Namely,

$$A = \int_{-\infty}^{\infty} p_r^{(0)}(z) \sin(\omega z)\, dz,$$

$$B = \int_{-\infty}^{\infty} \left(\frac{1}{r^{(0)}(z)^2} - \frac{1}{\bar{r}^2} \right) \cos\big[m\varphi^{(0)}(z)\big]\, dz.$$

The function M_1 has simple zeros for a discrete set of values of z_0 and φ_0. When the distance proportional to M_1 vanishes, this construction (along with the Melnikov theory) implies intersection of the stable and unstable manifolds of the whole manifold $\mathcal{M}_\varepsilon$, but it does *not* guarantee

that a_ε^s and a_ε^u lie on the stable and unstable manifolds of the *same* perturbed torus, T_ε. Thus, when M_1 vanishes along $\boldsymbol{n}_1$, heteroclinic orbits exist connecting (in general) *different* perturbed tori T_ε on $\mathcal{M}_\varepsilon$. In this case, the stable and unstable manifolds of the whole perturbed manifold $\mathcal{M}_\varepsilon$ do intersect, but the stable and unstable manifolds of a given perturbed torus (T_ε, a point on $\mathcal{M}_\varepsilon$) may *not* intersect. Hence, a transition chain may arise, leading to phase space transport, but no chaos will occur in this case, unless the stable and unstable manifolds of each perturbed torus T_ε also intersect. (This requires an additional condition.) As we discuss below, the mechanisms for phase space transport in this case include both Arnol'd diffusion (see, e.g. refs. [1, 5]), as well as lobe dynamics, since $W_\varepsilon^s(\mathcal{M}_\varepsilon)$ and $W_\varepsilon^u(\mathcal{M}_\varepsilon)$, being of codimension one, can separate regions of phase space into lobes.

Consider now the phase space connections for an individual perturbed torus T_ε, rather than the whole perturbed manifold $\mathcal{M}_\varepsilon$. Again we use the unperturbed homoclinic manifold $W(T_0)$ in order to coordinatize the manifolds $W_\varepsilon^s(T_\varepsilon)$ and $W_\varepsilon^u(T_\varepsilon)$. The manifold $W(T_0)$ is given by the two relations

$$H(r, p_r, p_\varphi) - \bar{H}(\bar{r}, 0, p_\varphi) = 0, \quad p_\varphi = \text{const.}$$

At each point a on $W(T_0)$, a plane Π is defined by the span of the two normal vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ at the point a. The homoclinic manifold $W(T_0)$ intersects the plane Π in precisely one point, namely the point a. Under perturbations, persistency of transversality implies that the intersections of Π with $W_\varepsilon^s(T_\varepsilon)$ and $W_\varepsilon^u(T_\varepsilon)$ will be discrete points. Again, one may choose the nearest intersection points in the sense of time of flight to T_ε. Note that these points, a_ε^s and a_ε^u, are now intersections in the plane Π of the stable and unstable orbits from the *same* perturbed torus, T_ε. Also, the vector $a_\varepsilon^s - a_\varepsilon^u$ now has *two* components (along the two normal directions $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ in the plane Π). Thus, to determine whether the manifolds W_ε^s and W_ε^u intersect we need to measure the distance between them along *both* directions. As before, the distance $a_\varepsilon^s - a_\varepsilon^u$ is measured to first order in ε by the Melnikov method. In this case, a zero of both components of the Melnikov vector $\boldsymbol{M} = (M_1, M_2)$ implies occurrence of not only transition chains, but also homoclinic intersections. The Melnikov vector components are given by

$$M_1 = \int_{-\infty}^{\infty} \boldsymbol{n}_1 \cdot \boldsymbol{g}\, dz$$

$$= \frac{\alpha}{2H_0^2} A \cos(\omega z_0) + \frac{m\beta p_\varphi}{2H_0^2} B \cos(m\varphi_0),$$

$$M_2 = \int_{-\infty}^{\infty} \boldsymbol{n}_2 \cdot \boldsymbol{g}\, dz = -\frac{m\beta}{2H_0} C \cos(m\varphi_0),$$

where

$$C = \int_{-\infty}^{\infty} \cos\left[m\varphi^{(0)}(z)\right] dz.$$

The component M_1 is computed as before. The integral for M_2, however, only converges conditionally. That is, sequences $\{z_j^+\}$ and $\{z_j^-\}$ with $z_j^+ \to \infty$ and $z_j^- \to -\infty$ can be chosen so that $\int_{z_j^-}^{z_j^+} \boldsymbol{n}_2 \cdot \boldsymbol{g}\, dz$ converges. (See refs. [8, 11].) This Melnikov vector $\boldsymbol{M} = (M_1, M_2)$ has simple zeros when both components vanish, i.e. when both $m\varphi_0 = \pm \frac{1}{2}\pi$ and $\omega z_0 = \pm \frac{1}{2}\pi$. In this case, we have both transition chains and homoclinic intersections.

At this point we may discuss the implications for chaos and phase space transport of the geometrical properties of these intersections. A given perturbed torus T_ε has three-dimensional stable and three-dimensional unstable manifolds. If both components of the Melnikov vector $\boldsymbol{M}$ vanish, these stable and unstable manifolds intersect in one dimension ($3 + 3 - 5 = 1$). This intersection is the orbit homoclinic to the perturbed torus. The three-dimensional stable manifold of T_ε intersects the four-dimensional unstable manifold of $\mathcal{M}_\varepsilon$ in two dimensions ($3 + 4 - 5 = 2$). So the majority of orbits on the stable manifold of T_ε do *not* come from $\mathcal{M}_\varepsilon$. The heteroclinic orbits, which do connect to T_ε from $\mathcal{M}_\varepsilon$, comprise the two-

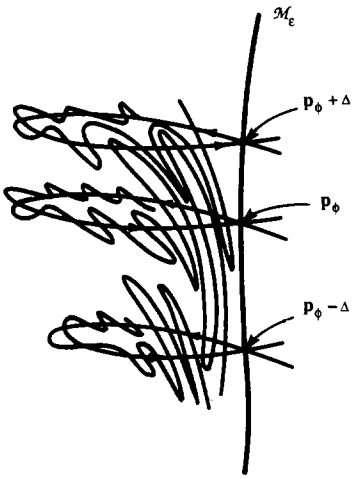

Fig. 7. Transition chain with both homoclinic and heteroclinic intersections.

dimensional intersection of the stable manifold of T_ε and the unstable manifold of $\mathcal{M}_\varepsilon$. Only one orbit on this two-dimensional intersection is homoclinic to T_ε. Thus, in this case a transition chain occurs (shown in fig. 7) and Arnol'd diffusion results, because almost all of the connecting orbits are from one perturbed torus to a different one. The presence of the homoclinic intersections may imply a horseshoe map construction from an iteration of the Poincaré map, but this has not been shown. Thus, homoclinic chaos in the presence of the combined perturbations cannot be ruled out, but also has not been proven.

Remark. Had only one of the Melnikov components vanished (in particular, the M_1 component), then the stable and unstable manifolds of $\mathcal{M}_\varepsilon$ would still have had three-dimensional intersections; and the stable manifold of a particular torus, T_ε, would still have intersected the unstable manifold of $\mathcal{M}_\varepsilon$ in two dimensions. However, with $M_2 \neq 0$ there would be no orbit homoclinic to the perturbed torus T_ε itself. (See fig. 8.) In this case, one might conclude there would be a transition chain and Arnol'd diffusion, but *without* chaos. (After all, without self-intersections, no horseshoe map could arise.) However, this situation is not generic. (See ref. [5], remark 4, p. 672.) Moreover, this situation does not arise here: since both of the conditions on φ_0 and z_0 above are fulfilled, the combined perturbations lead to vanishing of both components of the Melnikov vector.

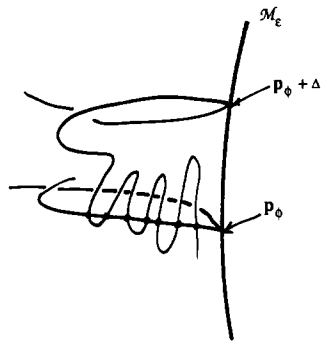

Fig. 8. Transition chain with only heteroclinic intersections.

Vanishing of either one, or both, of the Melnikov vector components, implies the presence of phase space transport by lobe dynamics, as discussed in ref. [13]. The basic requirement for lobe dynamics to dominate phase space transport is that the stable and unstable manifolds of the hyperbolic set are of codimension one, thereby allowing them to separate regions of phase space (in most situations). The hyperbolic set in the present case is $\mathcal{M}_\varepsilon$, whose four-dimensional stable and unstable manifolds are of codimension one in the five-dimensional phase space. Hence, lobe dynamics is the dominant mechanism for ray trapping and untrapping in fiber optics, when both axial and azimuthal perturbations are present. Lobe dynamics is also the dominant mechanism for ray trapping and untrapping in fiber optics, when only *one* of these perturbations is present.

Acknowledgements

This work was partially supported by the US Air Force Office of Scientific Research, contract number AFOSRISSA900024. The authors are also grateful to Tasso Kaper for helpful remarks.

References

[1] V.I. Arnol'd, Instability of dynamical systems with several degrees of freedom, Sov. Math. Dokl. 5 (1964) 581–585.

[2] M. Born and E. Wolf, Principles of Optics, Sixth (corrected) Ed. (Pergamon, Oxford, 1986).

[3] J. Guckenheimer and P.J. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields (Springer, New York, Heidelberg, Berlin, 1983).

[4] P.J. Holmes, Chaotic motions in a weakly nonlinear model for surface waves, J. Fluid Mech. 162 (1985) 365–388.

[5] P.J. Holmes and J.E. Marsden, Perturbations of n degree of freedom Hamiltonian systems with symmetry, J. Math. Phys. 23 (1982) 669–675.

[6] R.K. Luneberg, Mathematical Theory of Optics (University Microfilms, Ann Arbor, MI, 1944).

[7] V.K. Melnikov, On the stability of the center for time periodic perturbations, Trans. Moscow Math. Soc. 12 (1963) 1–57.

[8] C. Robinson, Horseshoes for autonomous Hamiltonian systems using the Melnikov integral, Ergod. Theory Dynam. Systems 8* (1988) 395–409.

[9] V. Rom-Kedar, A. Leonard and S. Wiggins, An analytical study of transport, mixing and chaos in an unsteady vortical flow, J. Fluid Mech. 214 (1990) 347–394.

[10] V. Rom-Kedar and S. Wiggins, Transport in two-dimensional maps, Arch. Rat. Mech. Anal. 109 (1990) 239–298.

[11] S. Wiggins, Global Bifurcations and Chaos: Analytical Methods (Springer, New York, Heidelberg, Berlin, 1988).

[12] S. Wiggins, Introduction to Applied Nonlinear Dynamical Systems and Chaos (Springer, New York, Heidelberg, Berlin, 1990).

[13] S. Wiggins, On the geometry of transport in phase space I. Transport in k-degree-of-freedom Hamiltonian systems, $2 \leq k \leq \infty$, Physica D 44 (1990) 471–501.

[14] K.B. Wolf, Lie Methods in Optics, Lecture Notes in Physics, Vol. 250 (Springer, New York, Heidelberg, Berlin, 1986).